# Basics:

for any experiment $\exists\ \Omega$ s.t. set of all possible sets.

Event: is a subset of $\Omega$, we say that an event occoures if outcome $\omega \in A \subset \Omega$ occours

* $\emptyset$ impossible event
* $\Omega$ certain event.

having $F \subset P(\Omega)$ is called event space.

$P(\Omega)$ is Power set of $\Omega := \{$all possible subsets$\}$

Properties:

$F \in \Omega$

$A \in F \to A^c \in F$

$A_i \in F \to \cup_i A_i \in F$

example coin toss

$\Omega = \{H, T\}$

event space $= \{\Omega, \emptyset, \{H\}, \{T\}\} = P(\Omega)$

dice throw $= \{\Omega, \emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$
all there complements.

if $|\Omega| < \infty$ then $F = P(\Omega)$

# Borel - $\sigma$ - Field

let $(\Omega, d)$ metric space, $\sigma$-Field can get created by open sets of $\Omega \Rightarrow$ Borel $\sigma$- field $\mathcal{B}(\Omega)$

\# $\mathcal{B}(\mathbb{R})$ contains all intervals $(a,b)$, $(a,b]$, $[a,b)$ $[a,b]$ $\forall a,b \in \mathbb{R}$, $a<b$, $(-\infty, a)$, $a \in \mathbb{R}$

$$\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R}) \Rightarrow \text{Powerset of } \mathbb{R}.$$

---

## Probability spaces:

let $\Omega$ sample space, $F$ is event space ( subset of $\Omega$

$P: F \rightarrow \mathbb{R}$ Probability measure is fullfiled

$P(A) \geq 0 \quad \forall A \in F$

$P(\Omega) = 1$

if $A_i \in F$, $i \in \mathbb{N}$ $A_i$ are disjoint $A_i \cap A_j = \emptyset$

$$P(A_i \cup A_j) = P(A_i) + P(A_j)$$

$$P(\bigcup A_i) = \sum P(A_i)$$

So $(\Omega, F, P)$ is probability space.

coin toss, $\Omega = \{H, T\}$ $\quad F = P(\Omega)$

$P : P(\Omega) \to R$ $\quad P(\emptyset) = 0$ $\quad P(\Omega) = 1$

$P(\{H\}) = \frac{1}{2}$ $\quad P(\{T\}) = \frac{1}{2}$

dice throw, $\Omega = \{1, 2, 3, 4, 5, 6\}$ $\quad P_c : P(\Omega) \to R$

$P_c(\emptyset) = 0$ $\quad P_c(\Omega) = 1$ $\quad P_c(\{1\}) = \frac{1}{6}$ - - - -

## Properties:

$P(A) \in [0, 1]$ $\quad P(\emptyset) = 0$

$P(A^c) = 1 - P(A)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A / B) = P(A) - P(A \cap B)$

$P(A) \leq P(B) \to A \subset B$

## Counting Rule:

given $(\Omega, P(\Omega), P)$ for $\forall A \subset P(\Omega)$

$$P(A) = \frac{|A|}{|\Omega|}$$

example $A = \{2, 3, 5\}$

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

## Random Variables:

**Measurable Functions:** Any function on finit spase is

let $\Omega, \Omega'$ and $F, F'$ (event spaces)

a function $X: \Omega \to \Omega'$ is called measurables

if $X^{-1}(A') \in F \quad \forall A' \in F'$

let $(\Omega, F, P)$ probability space, let $X: \Omega \to R$

$X$ is random variable, if $F$-$B(\Omega)$ measurable

**distribution of X:**

P.M $P_X: B(R) \to R$ $\quad P_X$ is $P \circ X^{-1}$

example

let $B \in B(R)$

$$P_X(B) = P(X^{-1}(B)) = P(\{w \in \Omega : X(w) \in B\}).$$

# for finite or countbly infinite $\Omega$ and $F = P(\Omega)$

every mapping $X : \Omega \to \mathbb{R}$ is Random variable.

## Random Variables

ⓐ Discret : distinct values
ⓑ Continous : any valus in an interval.

examples

$X = \begin{cases} 1 & \text{Head} \\ 0 & \text{Tail} \end{cases}$

Discret

(here we can't count)
$Y$ = mass of random animal
it takes interval.

Continous

$Z$ = year of student were born – discret.
1998 2000 2002 (countable)
even if it's infinit

$Q$ = number of ant born tmw (discret)

$X$ = exact winnig time to finish race so
this is continous random variable.

if we rounded the time it will be discret.

ex)) $X$ = number of heads after 3 filps.

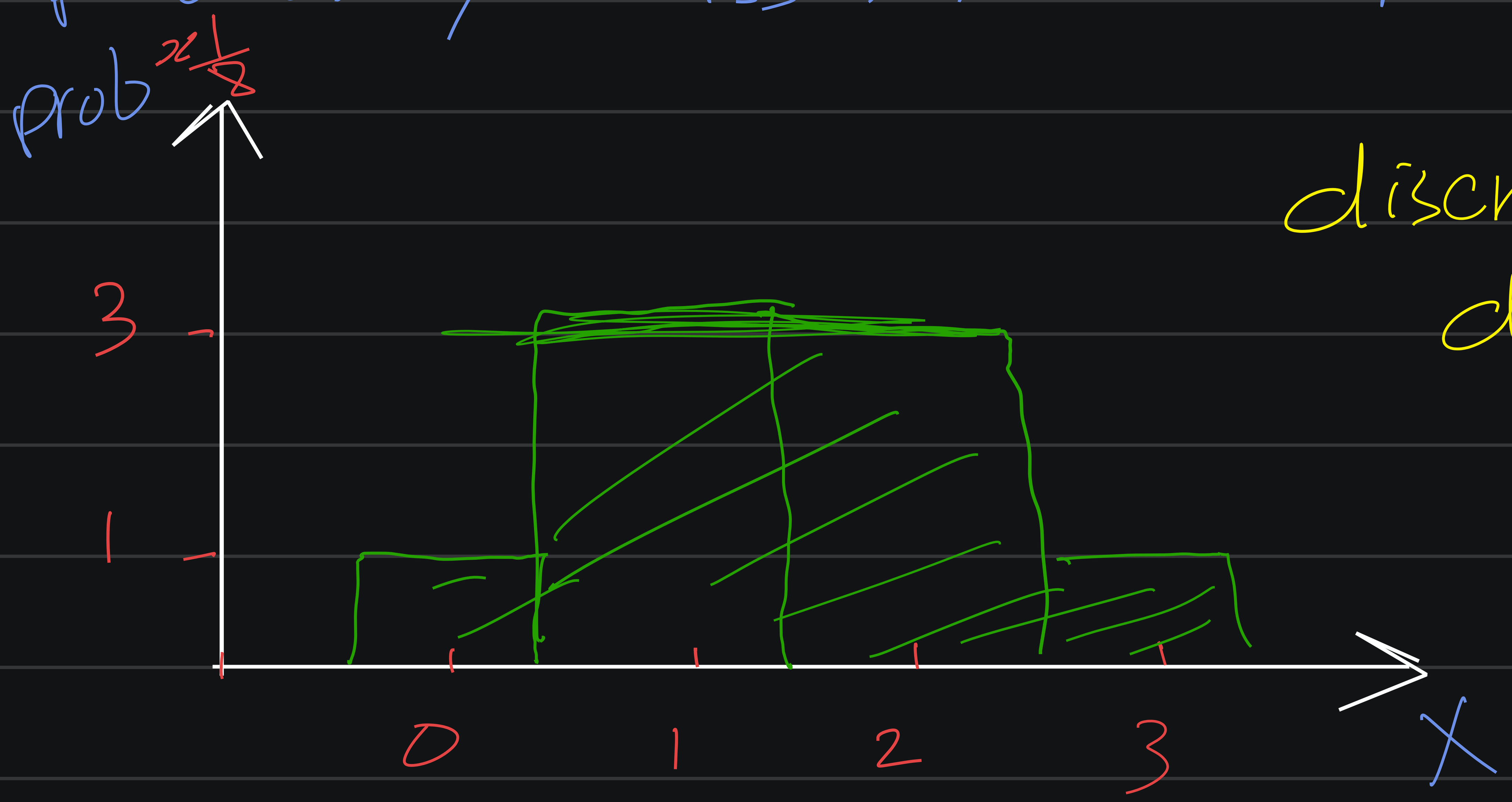
$P(X=0) = P(\{T,T,T\}) = \frac{1}{8}$

$P(X=1) = \frac{3}{8} = P(\{HTT, THT, TTH\})$

$P(X=2) = \frac{3}{8}$

$P(X=3) = \frac{1}{8}$

Probability mass function p.m.f




discret Probability distribution
p.m.f

$$\boxed{f_X(x_i) = P(X=x_i)}$$

in Bernolli experiment for $n$ times
so $X$ : number of head $X = \{0, 1, 2, \dots n\}$

$f_X(i) = \binom{n}{i} p^i (1-p)^{n-i}$ Binomal distribution

$f_X(2) = \binom{10}{2} \frac{1}{2}^2 \frac{1}{2}^8$

2 heads 10 filps

# Q/A Session:

Sample space = $\{1 \quad 1.3 \quad 1.7 \quad 2 \quad 2.3 \quad 2.7 \quad 3$
$3.3 \quad 3.7 \quad 4 \quad 4.3 \quad 4.7 \quad 5\}$

Possible events

$A_{best} = \{1.0\}$ $\quad A_{fail} = \{4, 4.3, 4.7, 5\}$

$A_{Pass} = \{1.0 \text{ ---- } 4.0\}$ subsets of sample space

$A_{good} = \{1.7, 2.0, 2.3\}$ all events = event space

$F = P(\Omega) = \{\emptyset, \Omega, \ldots\}$

Properties

$\Omega \in F$

$A \in F \longrightarrow A^c \in F$

$A_i \in F \Longrightarrow \bigcup_i A_i \in F$

if $\Omega$ is infinit
like $A = R$
we use Borel-σ-field:
we take (mid) all open sets $A \Rightarrow \sigma(A) = B(M)$

example $(R, |.|) \Longrightarrow B(R)$ = all open sets in $R$

distribution $(\Omega, F, P)$ $P: F \to \mathbb{R}$

$P_X := P \circ X^{-1}$

$X$ is random variable $X := \Omega \to \Omega'$

# if $\Omega$ is finit or countbly infinit $\to$ all $X: \Omega \to \mathbb{R}$ random variable

# Random Vector

it's just a vector of Random Variable

$(X_1, ---- X_n) := \Omega \longrightarrow R^n$ so $X$ (vector) is $F$-$B(R^n)$

$P_X = P_o X^{-1}$ we call it n-dimention random variable

so c.d.f

$$F_X(x) = P(X \leq x) = P(\{w \in \Omega \mid X_1(w) \leq x_1, ---- X_n(w) \leq x_n$$

example:

Tossing a coin two fine is dinomial distribution.

$\Omega = \{\{H, H\}, \{H, T\}, \{T, H\}, \{T, T\}\}$

$X$: number of heads $\quad Y$: number of tails
both are discret random variable.

if $P(H) = P \Rightarrow \frac{1}{2} \qquad P(T) = (1-P)$

$$F_{(X,Y)}(X=0, \underbrace{Y=2}) = \binom{2}{2} P^2 (1-P)^0$$
$$= 1 . (1-P)^2 P^0 = (1-P)^2$$
$$= \frac{1}{2}^2 = \boxed{\frac{1}{4}}$$

$$F_{X,Y}(X=1, Y=1) = \binom{2}{2} P^1 (1-P) +$$
$$\binom{2}{2} (1-P)^1 P^1$$
$$= 4 P(1-P) = 2 . (\tfrac{1}{2} \tfrac{1}{2})$$
$$= \boxed{\frac{1}{2}}$$

$F_{X,Y}(X=2,\ Y=0) = \binom{2}{2} p^2 (1-p)^0 = p^2 = \boxed{\frac{1}{4}}$

all other $p(X=x_i,\ Y=y_i)=0$ like $X=2$ $Y=2$ impossible.

---

## n-dimention Continous distribution

we say $X$ is n-dim cont dist if

$\exists$ non negative continous integrable function

$$f_{X_1 \cdots X_n} := \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$F_{X_1 \cdots X_n}(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1 \cdots X_n} \, dx_1 \cdots dx_n$$

this is n-dim density function

$$f_{X_1 \cdots X_n} = \frac{\partial F_{X_1 \cdots X_n}(x)}{\partial x_1 \cdots x_n}$$

---

So for vector r.v. we have $X \sim \mathcal{N}(\mu, C)$
$C \in \mathbb{R}^{n \times n}$ multi dimentional normal distribution

and marginal distribution it's the distribution of the $X$ in one direction only.

# Expectation of Random Variable:

if $X$ is discret r.v. P.m.f

$X$ number of heads 2 dice

if $X : \Omega \longrightarrow \{x_i\}$ lik $X_i : \Omega \rightarrow \{0, 1, 2\}$

so $\mu = E(X) = \sum_i x_i \, f_X(x_i)$

$E(X) = 0 \; P(X=0) + 1 \; P(X=1) + 2 \; P(X=2)$

$= 0 + \frac{1}{2} + 2 \cdot \frac{1}{4}$

$E(X)$ = expectation of the number of head $= 1$

---

if $X$ is continous: p.d.f

$$\mu = E(X) = \int_{-\infty}^{\infty} x \, f_X(x) \, . dx$$

this works also if there is a function

$g : R \rightarrow R$ (measurable). $g(X) \longrightarrow R$

so for discret $E(g(X)) = \sum_i g(x_i) \, f_X(x_i)$ p.m.f

Continous $E(g(X)) = \int_R g(x_i) \, f_X(x_i)$

p.d.f

## Properties:

- if $X \geq 0$ $\quad E(X) \geq 0$
- if $X = X_A$ for $A \in F$ $\Rightarrow$ $E(X) = P(A)$

describe only one even

- $E(a) = a$

$$E(a + bX) = E(a) + E(b)\, E(X)$$
$$a + bE(X)$$

even we have linear combination of random var.

## Variance / standard deviation:

let $(\Omega, F, P)$ with $X$ rand var with

$$\mu = E(X) \Rightarrow \sigma^2 = Var(X) = E\left((X-\mu)^2\right)$$

$$\sigma = \sqrt{Var(X)} = \sqrt{E\left((X-\mu)^2\right)}$$

for discret $Var(X) = \sum_i (x_i - \mu)^2 f_X(x_i)$

for continous $Var(X) = \int (x_i - \mu)^2 f_X(x)$

$$Var(X) = E(X^2) - E(X)^2$$

# Properties of Var

$$Var(X) = E(X^2) - E(X)^2$$

$$Var(a) = 0 \qquad Var(X) \geq 0$$

$$Var(a+bX) = b^2 Var(X)$$

$$Var(bX) = E(b^2X^2) - E(bX)^2$$
$$= b^2 E(X^2) - b^2 E(X)^2$$
$$= b^2 \left( E(X^2) - E(X)^2 \right)$$
$$Var(bX) = \boxed{b^2 Var(X)}$$

examples)

Binomial distribution:

$$f_X(x_i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$E(X) = np \qquad Var(X) = np(1-p)$$

X is Possion dist

$$f_X(x_i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

$$E(X) = \lambda \qquad Var(X) = \lambda$$

Uniform distribution

$$f_X(x_i) = \begin{cases} \frac{1}{b-a} & x \in (a,b) \\ 0 \end{cases}$$

$$E(X) = \frac{1}{2}(a+b) \qquad Var(X) = \frac{1}{12}(b-a)^2$$

Gaussian / normal dist

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$E(X) = \mu \qquad Var(X) = \sigma^2$$

## n-dimentional $E$ / Var

$$E(X_n) = (\mu_1 \dots \mu_n)^T$$

## Covariance and uncorrolated Random Variable

let $(\Omega, F, P)$ Probability space

$X = (X_1 \dots X_n)_i \quad i \in \{1, \dots n\}$

second moment exist $E(X_k X_l)$

so $\sigma_{kl} = Cov(X_k, X_l) = E((X_k - \mu_k)(X_l - \mu_l))$

so $C := Cov(X) := \sigma_{kl} \in R^{n \times n}$

$X_k, X_l$ is uncorrelated if $Cov(X_k X_l) = 0$

**Prop**

$C$ is Positive semi-defient

$Cov(X_k, X_k) = Var(X_k)$

$$C = E((X - \mu)(X - \mu)^T)$$

# Correlated coefficient

let $X = (Z, Y)^T$ rand vect

$Var(Z) > 0$ $Var(Y) > 0$ $Cov(X)$

Form $\leftarrow$ $\rho(Z, Y) = \frac{Cov(Z, Y)}{\sqrt{Var(Z) Var(Y)}}$

$\rho$ is correlated coeffient between $Z, Y$

## example

if $Z, Y$ are uncorrlated $Cov(Z, Y) = 0$

so $\rho(Z, Y) = 0$ not related

$\rho$ represent linear relationship between $Z, Y$

$\rho(Z, Y) \in [-1, 1]$

$\rho(Z, Y) = \pm 1 \iff P(Y = a + bZ) = 1$

iid : given $X = \{X_1 --- X_n\}$

we say that random variables of $X$ are called independent identical distribution

if $X_K$ is independent & $X \sim X_K$ $\forall K \in \{1, -- n\}$

---

2-D Gaussian distribution

let $(Z, Y)^T \sim \mathcal{N}(\mu, C)$ $\rightarrow$ $\mu = (\mu_z, \mu_y)^T$

$$C : Cov(Z, Y) = \begin{pmatrix} \sigma_z^2 & \rho\sigma_z\sigma_y \\ \rho\sigma_z\sigma_y & \sigma_y^2 \end{pmatrix}$$

if $Z, Y$ are independent then

$Z, Y$ are uncorlated $\rightarrow$ $Cov(Z, Y) = 0$

the oposite is false

if $(Z, Y)^T \sim \mathcal{N}(\mu, C)$

$(Z, Y)$ uncorrlated $\Leftrightarrow$ independent.